A multi-plane approach for ultrasound visual servoing : application to a registration task

Caroline Nadeau, Alexandre Krupa

A multi-plane approach for ultrasound visual servoing : application to a registration task

Caroline Nadeau and Alexandre Krupa

Abstract— **This paper presents a new image-based approach to control a robotic system equipped with an ultrasound imaging device. Moments based image features are extracted from three orthogonal ultrasound images to servo in-plane and out-of-plane motions of the system. Experimental results demonstrate that this approach improves upon techniques based on a single 2D US image in term of probe positioning. The second contribution of this paper is to use this method to perform a multimodal registration task by formulating it as a virtual visual servoing problem. Multimodal registration experiments performed with an ultrasound phantom containing an egg-shaped object provide a first experimental validation of the proposed method.**

Index Terms— **Ultrasound, visual servoing, multimodal registration**

I. INTRODUCTION

An increasing number of image-based robotic systems are developed to assist minimally invasive surgery procedures. Ultrasound (US) imaging devices are particularly well-adapted to such application insofar as they provide real time images during the operation. Moreover, in opposition to other modalities such as MRI or CT, US scanning is non invasive, low cost and may be repeated as often as necessary.

In this context, visual servoing approaches allow to directly control either the motion of the imaging device (eye-in-hand configuration) or the motion of a medical instrument (eye-to-hand configuration). In [1], the first application of an US based visual servoing was used to center the cross section of an artery in the image of the US probe during the tele-operation of this probe for diagnostic purpose. The in-plane motions of the probe were controlled by visual servoing while the other ones were tele-operated by the user. In [2], two degrees-of-freedom (DOF) of a needle-insertion robot are controlled by visual servoing to perform a percutaneous cholecystostomy while compensating involuntary patient motions. The target and the needle are automatically segmented in intraoperative US images and their respective poses, deduced from these data, are used to control the robot. However this control is once again limited to in-plane motions of the probe.

Some authors have proposed solutions to control out-of-plane motions of an US probe. In [3], a robotic system is proposed to track a surgical instrument and move it to a desired target. 3D US images are processed to localize respective positions of the target and the instrument tip. This position error is then used to control 4 DOF of the robotized tool in order to reach the target. However 3D US transducers provide currently volumes at a low acquisition rate which limits their use in real-time robotic applications. Another method using a 2D US probe is based on the model of the interaction between the object and the probe plane. In [4], two image points corresponding to the intersection of a surgical forceps with the image plane are used as visual features to control 4 DOF of the tool inside a beating heart. In relation with this work, the authors of [5] developed a predictive control scheme to keep the forceps visible in the US image.

More recently, a generalized US based servoing method was proposed to automatically reach a desired cross section of an organ of interest by servoing the 6 DOF of a medical robot holding a 2D US probe [7]. This method is based on the use of visual features built from image moments directly extracted from the US organ cross section. However, this method is a local approach since the convergence of the system is not guaranteed whatever the initial position. Moreover, symmetries on the organ geometry may lead to different probe positions that give a same cross section image of the organ.

In this paper, we present a new US image-based visual servoing approach used to control the 6 DOF of a robotic system equipped with a multi-plane US probe. The considered probe provides three 2D US images according to three orthogonal planes rigidly linked together (see Fig. 1). Therefore, we define in this paper a set of 2D features that can be extracted from these three planes and we model the corresponding interaction matrix. The second contribution of this paper is the application of the proposed control scheme to achieve a multimodal registration task, that we formulate as a virtual visual servoing approach. Image-to-image registration consists in finding the transformation between two image coordinates systems. These applications are particularly useful in medical field (a survey is presented in [6]) to superimpose the information provided by two different imaging modalities or to transfer a preoperative planning of a surgical gesture into the intraoperative field.

The structure of our paper is as follows. We initially describe the US image based control scheme and detail the benefits of the new features for a global convergence of the algorithm. We then present the considered application by expressing the registration task as a visual servoing problem. To validate our approach, servoing and registration results are presented and discussed in Section III and IV. Finally, concluding remarks and planned future works are given in Section V.

the authors are with IRISA, INRIA Rennes-Bretagne Atlantique, Campus de Beaulieu, 35042 Rennes cedex, France Caroline.Nadeau, Alexandre.Krupa@irisa.fr

II. ULTRASOUND VISUAL SERVOING

An image-based visual servoing control scheme consists in minimizing the error $\mathbf{e}(t) = \mathbf{s}(t) - \mathbf{s}^*$ between a current set of visual features $\mathbf{s}$ and a desired one $\mathbf{s}^*$. Considering an exponential decrease of this error, the classical control law [10] is given by:

$$\mathbf{v}_c = -\lambda \widehat{\mathbf{L}_\mathbf{s}}^{+} (\mathbf{s}(t) - \mathbf{s}^*), \quad (1)$$

where λ is the proportional gain involved in the exponential convergence of the error ($\dot{\mathbf{e}} = -\lambda\, \mathbf{e}$). In an eye-in-hand configuration, $\mathbf{v}_c$ is the instantaneous velocity applied to the visual sensor and $\widehat{\mathbf{L}_\mathbf{s}}^{+}$ is the pseudo-inverse of an estimation of the interaction matrix $\mathbf{L}_\mathbf{s}$ that relates the variation of the visual features to the velocity $\mathbf{v}_c$.

Traditional visual servoing control schemes refer to vision data acquired from a camera mounted on a robotic system. In this case, the vision sensor provides a projection of the 3D world to a 2D image and the coordinates of a set of 2D geometric primitives can be used to control the 6 DOF of the system. However, a 2D US transducer provides complete information in its image plane but not any outside of this plane. Therefore, US image based control laws can not rely only on features extracted from the 2D US scan to control the out-of-plane motions of the probe.

Only few previous works deal with the control of out-of-plane motions of a 2D US probe. Without a priori knowledge of the geometry of the object that interacts with the probe plane image, an US image based control scheme is proposed in [8] to servo the 6 DOF of an US probe in order to reach a desired image. Six geometric features are proposed to define the features vector $\mathbf{s}$. They represent the section of the object in the US plane by its mass center coordinates (x_g, y_g) and its main orientation α which are representative of the in-plane motions of the probe and present good decoupling properties. The area a of the object section, invariant to in-plane motions, and ϕ_1 and ϕ_2, moments invariant to the image scale, translation, and rotation are chosen to control out-of-plane motions. These features are computed from the image moments as follows:

$$\begin{cases} x_g = m_{10}/m_{00} \\ y_g = m_{01}/m_{00} \\ \alpha = \frac{1}{2}\, arctan(\frac{2\mu_{11}}{\mu_{20}-\mu_{02}}) \\ a = m_{00} \\ \phi_1 = \frac{{\mu_{11}}^2 - \mu_{20}\mu_{02}}{(\mu_{20}-\mu_{02})^2 + 4{\mu_{11}}^2} \\ \phi_2 = \frac{(\mu_{30}-3\mu_{12})^2 + (3\mu_{21}-\mu_{03})^2}{(\mu_{30}+\mu_{12})^2 + (\mu_{21}+\mu_{03})^2} \end{cases} \quad (2)$$

Where m_{ij} and μ_{ij} are respectively the moments and central moments of order $i+j$ that can be computed from the object contour C previously segmented in the image :

$$m_{ij} = \frac{-1}{j+1} \oint_C x^i\, y^{j+1}\, dx \quad (3)$$

The computation of the interaction matrix used to control in-plane and out-of-plane motions of the US probe is detailed in [8]. The time variation of moments of order $i+j$ is expressed in function of the probe velocity:

$$\dot{m}_{ij} = \mathbf{L}_{\mathbf{m_{ij}}} \mathbf{v}_c$$

with:

$$\mathbf{L}_{\mathbf{m_{ij}}} = [m_{v_x}\, m_{v_y}\, m_{v_z}\, m_{\omega_x}\, m_{\omega_y}\, m_{\omega_z}]$$

The components $(m_{v_x}, m_{v_y}, m_{\omega_z})$ related to the in-plane probe velocity are directly expressed from image moments. However the remaining components $(m_{v_z}, m_{\omega_x}, m_{\omega_y})$ also depend on the normal vector to the object surface which has to be estimated in each contour point. The final form of the resulting interaction matrix, given in [8] is:

$$\mathbf{L}_\mathbf{s} = \begin{bmatrix} -1 & 0 & x_{g_{v_z}} & x_{g_{\omega_x}} & x_{g_{\omega_y}} & y_g \\ 0 & -1 & y_{g_{v_z}} & y_{g_{\omega_x}} & y_{g_{\omega_y}} & -x_g \\ 0 & 0 & \alpha_{v_z} & \alpha_{\omega_x} & \alpha_{\omega_y} & -1 \\ 0 & 0 & \frac{a_{v_z}}{2\sqrt{a}} & \frac{a_{\omega_x}}{2\sqrt{a}} & \frac{a_{\omega_y}}{2\sqrt{a}} & 0 \\ 0 & 0 & \phi_{1_{v_z}} & \phi_{1_{\omega_x}} & \phi_{1_{\omega_y}} & 0 \\ 0 & 0 & \phi_{2_{v_z}} & \phi_{2_{\omega_x}} & \phi_{2_{\omega_y}} & 0 \end{bmatrix} \quad (4)$$

The efficiency of this control law highly depends on the object geometry and the pose of the initial image relative to the desired one. Indeed, in the case of symmetric objects, a given cross section of the object can be observed from multiple poses of the US probe. In this case, the information provided by one single US image is not sufficient to characterize the pose of the US probe relatively to the object. Therefore, the minimization of the image features error does not guarantee the global convergence of the algorithm in term of pose.

III. MULTI-PLANE ULTRASOUND VISUAL SERVOING APPROACH

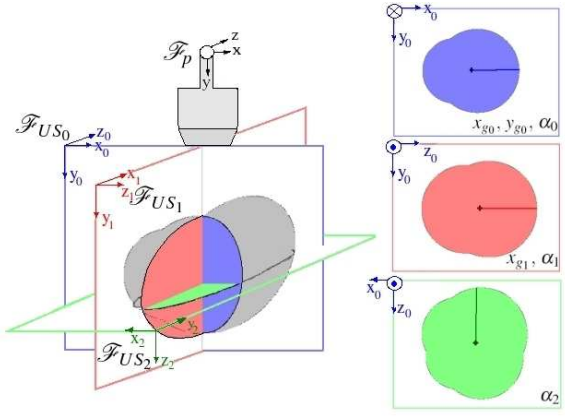

Fig. 1. The visual features are computed from three orthogonal planes. The probe frame coincides with the frame of US_0. On the right, this frame is reprojected in the various image plane frames

To overcome the local convergence limitation of the previous control scheme, we propose to consider a multi-plane probe p made up of 3 orthogonal planes (see Fig. 1). US_0 is aligned with the plane of the probe p and the plane US_1 (resp. US_2) corresponds to a rotation of 90° of this plane around the y_0-axis (resp. the x_0-axis). In such a configuration, we can note that each motion of the probe p corresponds to an in-plane motion in one of the three image planes. The in-plane velocities components (v_x, v_y, ω_z) of the probe

correspond to the in-plane motions $(v_{x_0}, v_{y_0}, \omega_{z_0})$ of the plane US_0, its out-of-plane components (v_z, ω_x) correspond to the in-plane velocities $(v_{x_1}, -\omega_{z_1})$ of the plane US_1 and finally its out-of-plane rotation velocity ω_y corresponds to the in-plane rotation velocity $-\omega_{z_2}$ of the plane US_2 (see Fig. 1).

Therefore, we propose to control the probe with six image features coupled to in-plane motions of the image plane where they are defined. More particularly, we will use the coordinates of the mass center of the object section, which are highly coupled to the in-plane translational motions and the orientation of the object section, which is representative of the in-plane rotation. The chosen image features vector is then:

$$\mathbf{s} = (x_{g_0}, y_{g_0}, x_{g_1}, \alpha_1, \alpha_2, \alpha_0). \tag{5}$$

A. Computation of the full interaction matrix

In each image plane US_i, the time variation of the moments-based image features (2) $\dot{\mathbf{s}}_i$ is related to the corresponding instantaneous velocity $\mathbf{v}_{c_i}$ through the interaction matrix (4):

$$\dot{\mathbf{s}}_i = \mathbf{L}_{\mathbf{s}_i}\, \mathbf{v}_{c_i}\, \forall i \in \{0,1,2\},$$

where the interaction matrix (4) can be written as:

$$\mathbf{L}_{\mathbf{s}_i} = [\mathbf{L}_{x_{g_i}}\, \mathbf{L}_{y_{g_i}}\, \mathbf{L}_{\alpha_i}\, \mathbf{L}_{A_i}\, \mathbf{L}_{l_{1_i}}]^T$$

In particular, each component of the features vector $\mathbf{s}$ detailed in (5) is related to the velocity of its corresponding image plane as follows:

$$\left\{ \begin{array}{rcl} \dot{x}_{g_0} & = & \mathbf{L}_{x_{g_0}} \mathbf{v}_{c_0} \\ \dot{y}_{g_0} & = & \mathbf{L}_{y_{g_0}} \mathbf{v}_{c_0} \\ \dot{x}_{g_1} & = & \mathbf{L}_{x_{g_1}} \mathbf{v}_{c_1} \\ \dot{\alpha}_1 & = & \mathbf{L}_{\alpha_1} \mathbf{v}_{c_1} \\ \dot{\alpha}_2 & = & \mathbf{L}_{\alpha_2} \mathbf{v}_{c_2} \\ \dot{\alpha}_0 & = & \mathbf{L}_{\alpha_0} \mathbf{v}_{c_0} \end{array} \right. \tag{6}$$

With the chosen configuration, the three planes frames are rigidly attached to the probe frame. We can therefore express the velocity $\mathbf{v}_{c_i}$ of each image plane in function of the instantaneous velocity of the probe $\mathbf{v}_c$:

$$\forall i \in \{0,1,2\},\ \mathbf{v}_{c_i} = {}^{\mathbf{US_i}}\mathbf{M_p}\, \mathbf{v}_c \tag{7}$$

with:

$${}^{\mathbf{US_i}}\mathbf{M_p} = \begin{pmatrix} {}^i\mathbf{R}_p & [{}^i\mathbf{t}_p]_\times\, {}^i\mathbf{R}_p \\ 0_3 & {}^i\mathbf{R}_p \end{pmatrix} \tag{8}$$

Where ${}^i\mathbf{t}_p$ and ${}^i\mathbf{R}_p$ are the translation vector and the rotation matrix of the probe frame $\mathscr{F}_p$ expressed in the coordinate system of the image plane $\mathscr{F}_{US_i}$.

Therefore, we can express the interaction matrix that relates the variation of the features vector (5) to the motion of the probe frame by:

$$\mathbf{L_s} = \begin{bmatrix} -1 & 0 & x_{g_{0_{v_z}}} & x_{g_{0_{\omega_x}}} & x_{g_{0_{\omega_y}}} & y_{g_0} \\ 0 & -1 & y_{g_{0_{v_z}}} & y_{g_{0_{\omega_x}}} & y_{g_{0_{\omega_y}}} & -x_{g_0} \\ x_{g_{1_{v_z}}} & 0 & -1 & y_{g_1} & x_{g_{1_{\omega_y}}} & x_{g_{1_{\omega_x}}} \\ \alpha_{1_{v_z}} & 0 & 0 & 1 & \alpha_{1_{\omega_y}} & \alpha_{1_{\omega_x}} \\ 0 & \alpha_{2_{v_z}} & 0 & \alpha_{2_{\omega_x}} & 1 & \alpha_{2_{\omega_y}} \\ 0 & 0 & \alpha_{0_{v_z}} & \alpha_{0_{\omega_x}} & \alpha_{0_{\omega_y}} & -1 \end{bmatrix} \tag{9}$$

B. The interaction matrix implemented in the control law

As stated previously, the six features chosen are coupled with one particular in-plane motion of their associated image plane. We propose then to relate their time variation only to the in-plane velocity components of their image frame. This means that we disregard the low variation of the image features due to the out-of-plane motions compared to the high variation due to in-plane motions.

The interaction matrix finally involved in the visual servoing control law (1) is then:

$$\widehat{\mathbf{L_s}} = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & y_{g_0} \\ 0 & -1 & 0 & 0 & 0 & -x_{g_0} \\ 0 & 0 & -1 & y_{g_1} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \tag{10}$$

Compared to the complete matrix given in (9), this one has great decoupling properties and is only dependent of the image features. In particular, the components of the estimated normal vector to the object surface are no longer involved. According to [10], the control scheme (1) is known to be locally asymptotically stable when a correct estimation $\widehat{\mathbf{L_s}}$ of $\mathbf{L_s}$ is used (i.e., as soon as $\mathbf{L_s}\widehat{\mathbf{L_s}}^{-1} > 0$).

C. Simulation validation

We compare the behavior of the control law based on features extracted from one cross section of the object or extracted from the three orthogonal images. We consider a mathematical object which is a compounding of four spheres of different radii. Considering this geometry of the object, the normal vector to its surface is perfectly known. Moreover, given a pose of the virtual probe the contour points of the object section are directly computed in the corresponding image.

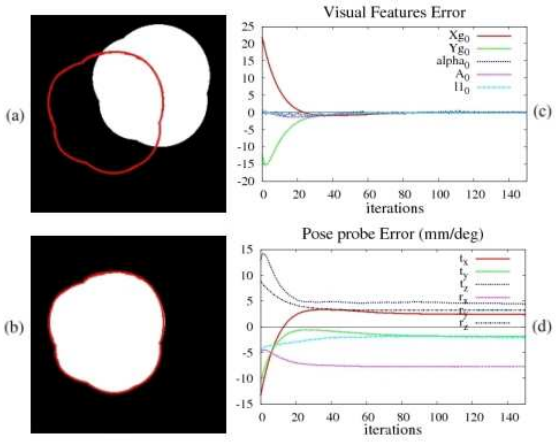

Fig. 2. (a), (b) Initial and final images of the probe. (c) Exponential decrease of the visual features errors. (d) Probe pose error in mm and deg, (the $\theta\mathbf{u}$ representation is used to describe the orientation).

By avoiding errors induced by the estimation of the normal vector or the detection of the object contour, we can efficiently compare both image-based algorithms. In Fig. 2,

a single image is considered to control the probe. During the convergence, the current section of the object (in white) and its desired contour (in red) are displayed. The expected exponential decrease of the error of the visual features is observed but the target pose is never reached because of the ambiguity of the object shape. On the other hand, by considering three orthogonal images the former ambiguity is resolved (see Fig. 3). In this case, the minimization of the visual features error leads to the desired pose. In both control laws a unitary gain λ is chosen and the computation time of one iteration of the algorithm is around $20ms$. The desired pose is then reached in $2s$ with the multi-plane control law.

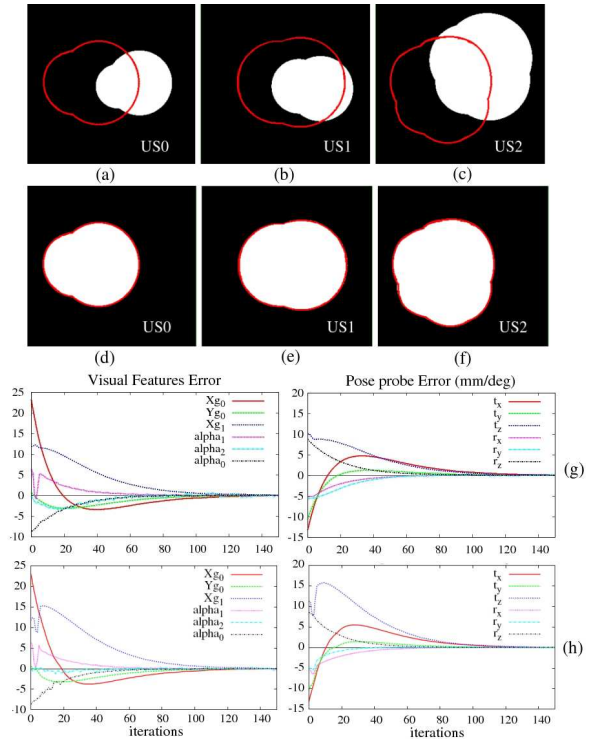

Fig. 3. (a), (b), (c) Images of the virtual probe at its initial pose. (d), (e), (f) Final images of the probe after the multi-plane algorithm convergence. Results obtained with the simplified interaction matrix (g) and the complete one (h) in term of image (left) and pose (right) error show a similar behavior of the control law and validate the simplification proposed.

The multi-plane approach overcomes the limitations of local convergence of the previous method. For servoing applications where a desired image has to be reached, its major limitation remains in the requirement of a specific imaging sensor to obtain the three orthogonal images of the object. However, this control law is well-adapted for other applications. We propose in the next section to apply this US image based control to perform a registration task with a classical 2D US probe.

IV. PRACTICAL APPLICATION TO A REGISTRATION TASK

Image-to-image registration methods are useful in medical field to transfer information from preoperative data to an intraoperative image. The aim is to compute the homogeneous transformation T_{reg} which transforms the coordinates of a pixel in the intraoperative image frame into voxel position expressed in the preoperative frame. Usual registration algorithms use an initialization of the parameters of this transformation based on artificial or anatomical landmarks identified in the preoperative 3D image and in a set of intraoperative US images. These parameters are then iteratively altered to optimize a similarity measure between both data sets according to a Powell-Brent search method. In our approach we propose to solve the registration task thanks to the previous image-based control scheme applied to a virtual multi-plane probe interacting with the preoperative volume.

A. Visual servoing formulation of the registration task

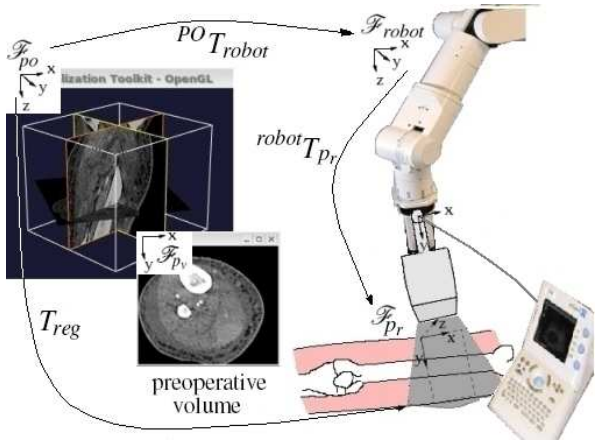

Fig. 4. On the left, a simulator is used to display the preoperative CT volume as well as a CT cross section corresponding to the current pose of the virtual probe. In parallel, the intraoperative image is acquired with an US probe mounted on a robotic arm.

The proposed system is detailed in Fig.4. An US probe, defined with the frame $\mathcal{F}_{p_r}$, is hold by a medical robot similar to the Hippocrate robot [9] and provides intraoperative images. In the robot base frame $\mathcal{F}_{robot}$, the probe pose $^{robot}\mathbf{P}_{p_r}$ is measured from the direct kinematic of the robotic arm. A 3D image of an organ is acquired preoperatively thanks to a medical imaging system.This preoperative volume expressed in frame $\mathcal{F}_{PO}$ is loaded in a software simulator that we have developed to reconstruct and display a dense volume with interpolation process from a set of parallel images. In addition to this display functionality, the simulator allows to model and control a virtual multi-plane probe, defined with the frame $\mathcal{F}_{p_v}$, interacting with the preoperative volume. For a given pose, this virtual probe generates three cross sections of the organ in the same imaging modality than the loaded volume. Image features extracted from this preoperative image are the current features of the control law while those extracted from the intraoperative US image are the desired ones. We then apply the multi-plane visual servoing control scheme to minimize the error between these current and desired features. After convergence of the algorithm, the pose of the virtual probe relative to the preoperative volume frame $^{PO}\mathbf{P}_{p_v}$ corresponds to the pose of the intraoperative image respectively to the preoperative one which characterizes the homogeneous transformation T_{reg}.

B. Practical setup

In practice, after the intraoperative US scan acquisition, a set of parallel images of the object is automatically acquired on both sides of this scan. A small intraoperative volume is then reconstructed and the two additional orthogonal images required for the multi-plane approach are created by a cubic interpolation process. In parallel, the virtual probe is arbitrarily positioned on the preoperative volume. The object contour is segmented with an active contour (snake) in the images provided by the virtual probe and the real one to compute the moments-based image features.

C. Registration results

Registration experiments are performed with an egg-shaped phantom (CIRS model 055) of size $3.9 \times 1.8 \times 1.8cm^3$. In the first application the preoperative volume is a 3D US volume, then a multimodal registration is performed with a 3D CT volume of the phantom. In both cases, the US intraoperative images are acquired using a Sonosite 180 2D US machine connected to a convex 2-5MHz US transducer with a depth of $12cm$. The US images resolution is 768×576 with a pixel size of $0.3 \times 0.3mm^2$, while the CT images resolution is 512×512 with a pixel size of $0.45 \times 0.45mm^2$.

1) US-US registration: The preoperative volume loaded in the simulator is created from a set of 250 parallel images acquired every $0.25mm$ during a translational motion of the probe. For validation purpose, the first image of this sequence is used to compute the transformation ${}^{PO}T_{robot}$. Indeed its pose is known in the preoperative frame ${}^{PO}\mathbf{P}_i$ such as in the robotic frame ${}^{robot}\mathbf{P}_i$ thanks to the direct kinematics of the robotic arm. Then, without moving the phantom, we position the probe on a new location which is considered as the intraoperative one. The corresponding image of the 2D probe is considered as the first intraoperative scan. Then a small translational motion is applied to this probe to acquire a set of parallel images from which the additional orthogonal images can be extracted by interpolation process.

In the preoperative US volume a virtual multi-plane probe is arbitrarily initially positioned then controlled as described in Section II to automatically perform the registration. The results are presented in Fig. 5. In this experiment we maintain unchanged the position of the physical object between preoperative and intraoperative images acquisitions to obtain a ground truth of the registration transformation T_{reg} thanks to the robot odometry:

$$T_{reg} = {}^{PO}T_{robot}\ {}^{robot}T_{p_r}$$

To validate our approach in term of probe positioning, the desired pose of the virtual probe in the preoperative frame ${}^{PO}\mathbf{P}^*_{p_v}$ is computed in the following way:

$${}^{PO}\mathbf{P}^*_{p_v} = {}^{PO}T_{robot}\ {}^{robot}\mathbf{P}_{p_r}$$

where ${}^{robot}\mathbf{P}_{p_r}$ is the pose of the real probe in the intraoperative frame, given by the robot odometry. The convergence of the algorithm in term of pose is then assessed by computing the error between the current and the desired pose of the virtual probe both expressed in the preoperative frame.

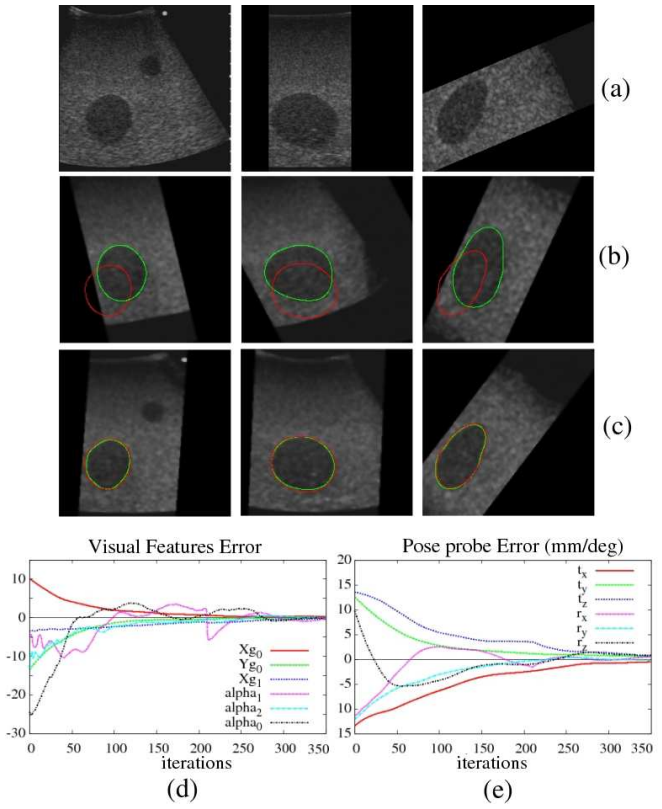

Fig. 5. (a) Intraoperative scan (on left) and interpolated additional orthogonal images (on right). (b), (c) Initial and final images of the virtual probe in the preoperative volume frame. (d), (e) Convergence of the algorithm in term of visual features and pose. A gain $\lambda = 0.8$ is used in the control law. With an iteration time of $40ms$, the registration task is performed in $12s$.

Five additional tests are run from different initial poses. The results are presented in the table I. The global convergence of the method is assessed by choosing large initial errors on the registration parameters. For each initial pose (1 to 5 in the table), the error of the virtual probe pose expressed in the preoperative frame is given in translation and rotation before and after the image features convergence. Initial translation errors up to $2.2cm$, which is more than the half-size of the object, and rotation errors up to 90° have been used in these tests. The mean error on the registration transformation is $1.92mm$ with a standard deviation of $0.83mm$ in translation and 2.90° with a standard deviation of 1.36° in rotation. Despite the symmetry of the object along its main axis, the pose convergence of the algorithm is therefore efficiently obtained with the multi-plane approach.

TABLE I

INITIAL AND FINAL POSE ERRORS OF THE VIRTUAL PROBE

Pose Error		1		2		3		4		5	
		init	final	init	final	init	final	init	final	init	final
T (mm)	tx	-13.25	1.13	-20.5	0.006	-13.47	-1.20	4.7	-0.3	-6.5	-0.4
	ty	-1.5	-0.02	22.6	-0.4	12.6	0.3	13.2	0.2	20.8	0.1
	tz	14.2	-0.4	20.5	-1.3	13.5	0.8	-2.5	-2.4	-14.5	-3.1
R(°)	Rx	-11.5	0.5	8.0	0.3	-11.5	1.0	89.4	4.0	75.6	4.0
	Ry	-6.3	0.26	12.0	0.17	-12.0	-1.8	1.4	1.0	-1.0	-0.6
	Rz	9.7	2.9	10.3	0.86	9.7	-1.0	1.2	-0.8	38.4	-0.8

2) US-CT multimodal registration: Intraoperative images are acquired as described previously and in the preoperative CT volume, the initial pose of the virtual probe is arbitrarily chosen. The only requirement is to visualize the entire cross section of the object in the three images of the multi-plane probe. The desired features are computed from the US intraoperative images scaled to fit the CT pixel resolution (see Fig. 6). The corresponding desired contours (in red) are extracted and reprojected in the CT images with the current contours (in green).

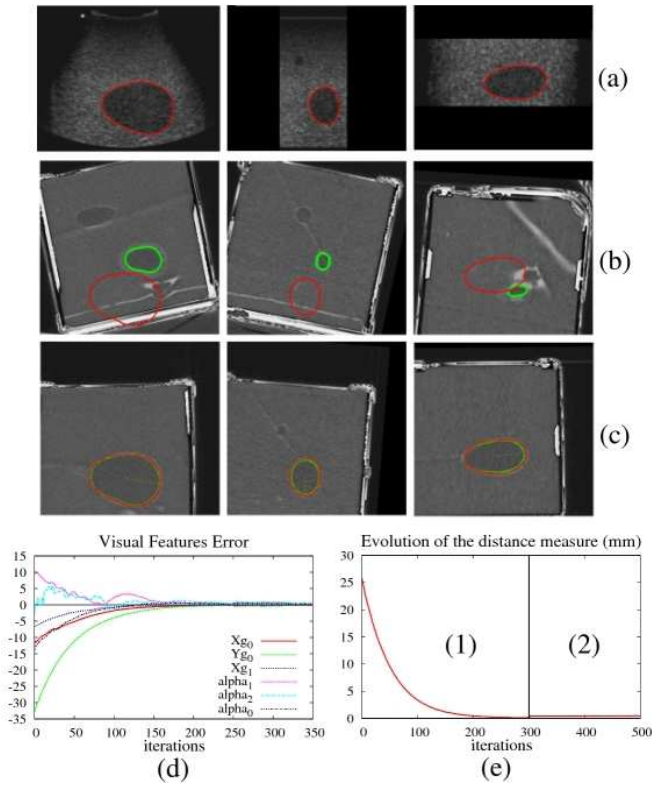

Fig. 6. (a) Intraoperative scan (on left) and interpolated additional orthogonal images (on right). (b), (c) Initial and final preoperative CT images. (d) Evolution of the features error. A gain $\lambda = 0.8$ is used in the control law and the features convergence is performed in 8 s. (e) Position error of the object section mass center during the visual servoing algorithm (part (1)) and an open loop motion (part (2)).

The exponential decrease of the features error is observed, however there is this time no ground truth to estimate the positioning error of the virtual probe. Therefore, to assess the precision of the registration, we apply after the algorithm convergence, an open-loop translational motion of $2cm$ along the z-axis to the real and virtual probes. During this motion, two corresponding sequences of intraoperative and preoperative images are acquired in which the coordinates of the object section mass center are extracted. We then quantify the misalignment between the preoperative and intraoperative images by using the following distance error:

$$d = \sqrt{({}^{US}x_{g_0} - {}^{CT}x_{g_0})^2 + ({}^{US}y_{g_0} - {}^{CT}y_{g_0})^2}$$

During the servoing process, the distance error is minimized from 39.3 to $0.27mm$. Then during the open-loop motion, the distance error ranges from 0.3 to $0.7mm$ (see Fig. 6(e)), which demonstrates the accuracy of the registration task.

V. CONCLUSIONS

This paper presented a new method of US visual servoing based on image moments to control an US device mounted on a robotic arm in an eye-in-hand configuration. We designed six features extracted from three orthogonal image planes to efficiently control in-plane and out-of-plane motions of the system. In particular, we applied our visual servoing control law to deal with the image registration problem which consists in computing the transformation between two image frames. In medical field, the purpose is to match an intraoperative US image with a preoperative volume of an organ in order to transfer preoperative information into the intraoperative field. The problem is here solved by considering a virtual probe attached to the preoperative volume. The multi-plane control scheme is then applied to this probe which is moved until that its intersection with the volume corresponds to the intraoperative image.

Further work will be undertaken to address the issue of physiological motions and non rigid registration. The challenge remains also in the US images processing to deal with cases where the organ can not be segmented with a closed contour.

ACKNOWLEDGMENT

The authors acknowledge the support of the ANR project PROSIT of the French National Research Agency.

REFERENCES

[1] P. Abolmaesumi, M. Sirouspour, S. Salcudean, W. Zhu, Adaptive image servo controller for robot-assisted diagnostic ultrasound,*Proceedings of IEEE/ASME Int. Conf. on Advanced Intelligent Mechatronics*, Vol. 2, pp. 1199-1204, Como, Italy, July 2001.

[2] J. Hong, T. Dohi, M. Hashizume, K. Konishi, N. Hata, A motion adaptable needle placement instrument based on tumor specific ultrasonic image segmentation. *In 5th Int. Conf. on Medical Image Computing and Computer Assisted Intervention, MICCAI'02*, pp122-129, Tokyo, Japan, September 2002.

[3] J.A. Stoll, P.M. Novotny, R.D. Howe and P.E. Dupont, Real-time 3D Ultrasound-based Servoing of a Surgical Instrument. *In IEEE Int. Conf. on Robotics and Automation, ICRA'06*, pp. 613-618, Orlando, Florida, USA, May 2006.

[4] M.A. Vitrani, H. Mitterhofer, N. Bonnet, G. Morel, Robust ultrasound-based visual servoing for beating heart intracardiac surgery. *In IEEE Int. Conf. on Robotics and Automation, ICRA'07*, pp3021-3027, Roma, Italy, April 2007.

[5] M. Sauvee, P. Poignet, E. Dombre, US image based visual servoing of a surgical instrument through non-linear model predictive control, *Int. Journal of Robotics Research*, vol. 27, no. 1, January 2008.

[6] J. P. W. Pluim, J. B. A. Maintz, M. A. Viergever, Mutual Information Based Registration of Medical Images: A Survey, *In IEEE Trans. Med. Imaging*, vol. 22, pp. 986-1004, August 2003.

[7] R.Mebarki, A. Krupa and F. Chaumette, Image moments-based ultrasound visual servoing. *In IEEE Int. Conf. on Robotics and Automation, ICRA'08*, pp 113-119, Pasadena, CA, USA, May 2008.

[8] R.Mebarki, A. Krupa and F. Chaumette, 2D ultrasound probe complete guidance by visual servoing using image moments. *In IEEE Trans. on Robotics*, vol. 26, nr. 2, pp 296-306; 2010.

[9] F. Pierrot, E. Dombre, E. Degoulange, L. Urbain, P. Caron, S. Boudet, J. Gariepy and J. Megnien, Hippocrate: A safe robot arm for medical applications with force feedback, *In Medical Image Analysis (MedIA)*, vol. 3, no. 3, pp 285-300; 1999.

[10] B. Espiau, F. Chaumette and P. Rives, A new approach to visual servoing in robotics. *In IEEE Trans. on Robotics*, 8(3):313-326, 1992.